

# 7 OIPE

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/554,941

TIME: 16:24:51

Input Set : A:\SUBSTITUTE SEQUENCE LISTING.txt

Output Set: N:\CRF3\09072001\I554941.raw

## SEQUENCE LISTING

ENTERED

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Atkinson, Howard J

7 McPherson, Michael J

8 Urwin, Peter E

10 (ii) TITLE OF INVENTION: Proteinase Inhibitor Fusion Proteins

12 (iii) NUMBER OF SEQUENCES: 21

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Syngenta Biotechnology, Inc.

16 (B) STREET: 3054 Cornwallis Road

17 (C) CITY: RTP

18 (D) STATE: NC

19 (E) COUNTRY: USA

20 (F) ZIP: 27709

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Floppy disk

24 (B) COMPUTER: IBM PC compatible

25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/554,941

C--> 30 (B) FILING DATE: 22-May-2000

31 (C) CLASSIFICATION:

33 (viii) ATTORNEY/AGENT INFORMATION:

34 (A) NAME: Meigs, J. Timothy

35 (B) REGISTRATION NUMBER: 38,241

36 (C) REFERENCE/DOCKET NUMBER: S-30287A

38 (ix) TELECOMMUNICATION INFORMATION:

39 (A) TELEPHONE: 919-541-8587

42 (2) INFORMATION FOR SEQ ID NO: 1:

44 (i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 14 amino acids

46 (B) TYPE: amino acid

47 (C) STRANDEDNESS: single

48 (D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: peptide

52 (iii) HYPOTHETICAL: NO

C--> 54 (iv) ANTI-SENSE: NO

56 (v) FRAGMENT TYPE: internal

58 (vi) ORIGINAL SOURCE:

59 (A) ORGANISM: PsMTa Linker

62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

64 Val Ile Leu Gly Val Gly Pro Ala Lys Ile Gln Phe Glu Gly

65 1 5 10

67 (2) INFORMATION FOR SEQ ID NO: 2:

69 (i) SEQUENCE CHARACTERISTICS:

70 (A) LENGTH: 11 amino acids

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/554,941

DATE: 09/07/2001

TIME: 16:24:51

Input Set : A:\SUBSTITUTE SEQUENCE LISTING.txt

Output Set: N:\CRF3\09072001\I554941.raw

```

71          (B) TYPE: amino acid
72          (C) STRANDEDNESS: single
73          (D) TOPOLOGY: linear
75      (ii) MOLECULE TYPE: peptide
77      (iii) HYPOTHETICAL: NO
C--> 79      (iv) ANTI-SENSE: NO
81          (v) FRAGMENT TYPE: internal
83          (vi) ORIGINAL SOURCE:
84              (A) ORGANISM: galactose oxidase linker
87          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
89      Gln Ala Ser Ser Tyr Thr Ala Pro Gln Pro Gln
90      1          5          10
92 (2) INFORMATION FOR SEQ ID NO: 3:
94      (i) SEQUENCE CHARACTERISTICS:
95          (A) LENGTH: 30 base pairs
96          (B) TYPE: nucleic acid
97          (C) STRANDEDNESS: single
98          (D) TOPOLOGY: linear
100      (ii) MOLECULE TYPE: DNA (genomic)
102      (iii) HYPOTHETICAL: NO
C--> 104      (iv) ANTI-SENSE: NO
106          (vi) ORIGINAL SOURCE:
107              (A) ORGANISM: oligo P1
110          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
112 ATGTCGAGCG ACGGACGGCC GGTGCTTGGC
114 (2) INFORMATION FOR SEQ ID NO: 4:
116      (i) SEQUENCE CHARACTERISTICS:
117          (A) LENGTH: 48 base pairs
118          (B) TYPE: nucleic acid
119          (C) STRANDEDNESS: single
120          (D) TOPOLOGY: linear
122      (ii) MOLECULE TYPE: DNA (genomic)
124      (iii) HYPOTHETICAL: NO
C--> 126      (iv) ANTI-SENSE: NO
128          (vi) ORIGINAL SOURCE:
129              (A) ORGANISM: oligo P2
132          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
134 GATCTTCGCC GGACCGACGC CAAGAATCAC GGCATTGCA CTGGCATC
136 (2) INFORMATION FOR SEQ ID NO: 5:
138      (i) SEQUENCE CHARACTERISTICS:
139          (A) LENGTH: 48 base pairs
140          (B) TYPE: nucleic acid
141          (C) STRANDEDNESS: single
142          (D) TOPOLOGY: linear
144      (ii) MOLECULE TYPE: DNA (genomic)
146      (iii) HYPOTHETICAL: NO
C--> 148      (iv) ANTI-SENSE: NO
150          (vi) ORIGINAL SOURCE:
151              (A) ORGANISM: oligo P3

```

30

48

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/554,941

TIME: 16:24:51

Input Set : A:\SUBSTITUTE SEQUENCE LISTING.txt

Output Set: N:\CRF3\09072001\I554941.raw

```

154      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
156 GTCGGTCCGG CGAAGATCCA GTTTGAAGGT AGTAATCATC ATGATGAC
158 (2) INFORMATION FOR SEQ ID NO: 6:
160      (i) SEQUENCE CHARACTERISTICS:
161          (A) LENGTH: 32 base pairs
162          (B) TYPE: nucleic acid
163          (C) STRANDEDNESS: single
164          (D) TOPOLOGY: linear
166      (ii) MOLECULE TYPE: DNA (genomic)
168      (iii) HYPOTHETICAL: NO
C--> 170      (iv) ANTI-SENSE: NO
172      (vi) ORIGINAL SOURCE:
173          (A) ORGANISM: oligo P4
176      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
178 TTCTTACTCA TCATCTTCAT CCCTGGACTT GC
180 (2) INFORMATION FOR SEQ ID NO: 7:
182      (i) SEQUENCE CHARACTERISTICS:
183          (A) LENGTH: 45 base pairs
184          (B) TYPE: nucleic acid
185          (C) STRANDEDNESS: single
186          (D) TOPOLOGY: linear
188      (ii) MOLECULE TYPE: DNA (genomic)
190      (iii) HYPOTHETICAL: NO
C--> 192      (iv) ANTI-SENSE: NO
194      (vi) ORIGINAL SOURCE:
195          (A) ORGANISM: oligo P5
198      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
200 CTGGGGGGGCT GTGTAAGAAC TAGCTTGGGC ATTTGCACTG GCATC
202 (2) INFORMATION FOR SEQ ID NO: 8:
204      (i) SEQUENCE CHARACTERISTICS:
205          (A) LENGTH: 45 base pairs
206          (B) TYPE: nucleic acid
207          (C) STRANDEDNESS: single
208          (D) TOPOLOGY: linear
210      (ii) MOLECULE TYPE: DNA (genomic)
212      (iii) HYPOTHETICAL: NO
C--> 214      (iv) ANTI-SENSE: NO
216      (vi) ORIGINAL SOURCE:
217          (A) ORGANISM: oligo P6
220      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
222 AGTTCTTACA CAGCCCCCA GCCTGGTAGT AATCATCATG ATGAC
224 (2) INFORMATION FOR SEQ ID NO: 9:
226      (i) SEQUENCE CHARACTERISTICS:
227          (A) LENGTH: 31 base pairs
228          (B) TYPE: nucleic acid
229          (C) STRANDEDNESS: single
230          (D) TOPOLOGY: linear
232      (ii) MOLECULE TYPE: DNA (genomic)
234      (iii) HYPOTHETICAL: NO

```

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/554,941

TIME: 16:24:51

Input Set : A:\SUBSTITUTE SEQUENCE LISTING.txt

Output Set: N:\CRF3\09072001\I554941.raw

```

C--> 236      (iv) ANTI-SENSE: NO
      240      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
      242 ACTATGGATC CAGTAATCAT CATGATGACT C                               31
      244 (2) INFORMATION FOR SEQ ID NO: 10:
      246      (i) SEQUENCE CHARACTERISTICS:
      247          (A) LENGTH: 29 base pairs
      248          (B) TYPE: nucleic acid
      249          (C) STRANDEDNESS: single
      250          (D) TOPOLOGY: linear
      252      (ii) MOLECULE TYPE: DNA (genomic)
      254      (iii) HYPOTHETICAL: NO
C--> 256      (iv) ANTI-SENSE: NO
      260      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
      262 ATATTAAGCT TTTCTTACTC ATCATCTTC                               29
      264 (2) INFORMATION FOR SEQ ID NO: 11:
      266      (i) SEQUENCE CHARACTERISTICS:
      267          (A) LENGTH: 14 amino acids
      268          (B) TYPE: amino acid
      269          (C) STRANDEDNESS: single
      270          (D) TOPOLOGY: linear
      272      (ii) MOLECULE TYPE: peptide
      274      (iii) HYPOTHETICAL: NO
C--> 276      (iv) ANTI-SENSE: NO
      278      (v) FRAGMENT TYPE: internal
      280      (vi) ORIGINAL SOURCE:
      281          (A) ORGANISM: modified galactose oxidase linker
      284      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
      286      Gln Ala Ser Ile Glu Gly Arg Tyr Thr Ala Pro Gln Pro Gln
      287      1          5          10
      289 (2) INFORMATION FOR SEQ ID NO: 12:
      291      (i) SEQUENCE CHARACTERISTICS:
      292          (A) LENGTH: 18 amino acids
      293          (B) TYPE: amino acid
      294          (C) STRANDEDNESS: single
      295          (D) TOPOLOGY: linear
      297      (ii) MOLECULE TYPE: peptide
      299      (iii) HYPOTHETICAL: NO
C--> 301      (iv) ANTI-SENSE: NO
      303      (v) FRAGMENT TYPE: internal
      305      (vi) ORIGINAL SOURCE:
      306          (A) ORGANISM: natural immunoglobulin linker
      309      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
      311      Lys Glu Ser Gly Ser Val Ser Ser Glu Gln Leu Ala Gln Phe Arg Ser
      312      1          5          10          15
      314      Leu Asp
      317 (2) INFORMATION FOR SEQ ID NO: 13:
      319      (i) SEQUENCE CHARACTERISTICS:
      320          (A) LENGTH: 13 amino acids
      321          (B) TYPE: amino acid

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/554,941

DATE: 09/07/2001

TIME: 16:24:51

Input Set : A:\SUBSTITUTE SEQUENCE LISTING.txt

Output Set: N:\CRF3\09072001\I554941.raw

```

322          (C) STRANDEDNESS: single
323          (D) TOPOLOGY: linear
325      (ii) MOLECULE TYPE: peptide
327      (iii) HYPOTHETICAL: NO
C--> 329      (iv) ANTI-SENSE: NO
331          (v) FRAGMENT TYPE: internal
333          (vi) ORIGINAL SOURCE:
334              (A) ORGANISM: linker peptide
337      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
339      Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Pro
340          1             5             10
342 (2) INFORMATION FOR SEQ ID NO: 14:
344      (i) SEQUENCE CHARACTERISTICS:
345          (A) LENGTH: 15 amino acids
346          (B) TYPE: amino acid
347          (C) STRANDEDNESS: single
348          (D) TOPOLOGY: linear
350      (ii) MOLECULE TYPE: peptide
352      (iii) HYPOTHETICAL: NO
C--> 354      (iv) ANTI-SENSE: NO
356          (v) FRAGMENT TYPE: internal
358          (vi) ORIGINAL SOURCE:
359              (A) ORGANISM: linker peptide
362      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
364      Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
365          1             5             10             15
367 (2) INFORMATION FOR SEQ ID NO: 15:
369      (i) SEQUENCE CHARACTERISTICS:
370          (A) LENGTH: 31 base pairs
371          (B) TYPE: nucleic acid
372          (C) STRANDEDNESS: single
373          (D) TOPOLOGY: linear
375      (ii) MOLECULE TYPE: DNA (genomic)
377      (iii) HYPOTHETICAL: NO
C--> 379      (iv) ANTI-SENSE: NO
381          (vi) ORIGINAL SOURCE:
382              (A) ORGANISM: oligo P7
385      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
387 ACATGTCGAA TTCTTAGGCA TTTGCACTGG C
389 (2) INFORMATION FOR SEQ ID NO: 16:
391      (i) SEQUENCE CHARACTERISTICS:
392          (A) LENGTH: 24 base pairs
393          (B) TYPE: nucleic acid
394          (C) STRANDEDNESS: single
395          (D) TOPOLOGY: linear
397      (ii) MOLECULE TYPE: DNA (genomic)
399      (iii) HYPOTHETICAL: NO
C--> 401      (iv) ANTI-SENSE: NO
403          (vi) ORIGINAL SOURCE:

```

31

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/554,941

DATE: 09/07/2001

TIME: 16:24:52

Input Set : A:\SUBSTITUTE SEQUENCE LISTING.txt

Output Set: N:\CRF3\09072001\I554941.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:54 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:79 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:104 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:126 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:148 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:170 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:192 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:214 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:236 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:256 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:276 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:301 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:329 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:354 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:379 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:401 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:423 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:445 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:467 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:489 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]  
L:509 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]